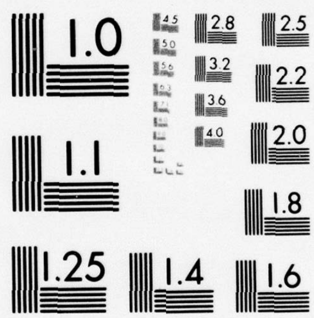
AD-A037 008 OHIO STATE UNIV COLUMBUS DEPT OF GEODETIC SCIENCE F/G 8/5
LEAST-SQUARES COLLOCATION AS A GRAVITATIONAL INVERSE PROBLEM.(U)
NOV 76 H MORITZ F19628-76-C-0010
UNCLASSIFIED DGS-249 AFGL-TR-76-0278 NL

1 OF 1
ADA037008

END
DATE
FILMED
4-77



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

AFGL-TR-76-0278

LEAST-SQUARES COLLOCATION AS A GRAVITATIONAL INVERSE PROBLEM

Helmut Moritz

The Ohio State University
Research Foundation
Columbus, Ohio 43212

November 1976

Scientific Report No. 6

AIR FORCE GEOPHYSICS LABORATORY
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
HANSCOM AFB, MASSACHUSETTS 01731

Qualified requestors may obtain additional copies from the Defense Documentation Center. All others should apply to the National Technical Information Service.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

19 REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 18 AFGL-TR-76-0278

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER: 9 Interim rept.

4. TITLE (and Subtitle): 6 Least-squares Collocation as a Gravitational Inverse Problem.

5. TYPE OF REPORT & PERIOD COVERED: Scientific. Interim. Scientific Report No. 6

6. PERFORMING ORG. REPORT NUMBER: Report No. 249

7. AUTHOR(s): 10 Helmut Moritz

8. CONTRACT OR GRANT NUMBER(s): 15 F19628-76-C-0010

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
Department of Geodetic Science
The Ohio State University - 1958 Neil Avenue
Columbus, Ohio 43210

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
62101F
76000302

11. CONTROLLING OFFICE NAME AND ADDRESS:
Contract Monitor: Bela Szabo/LW
Air Force Geophysics Laboratory
Hanscom AFB, Massachusetts 01731

12. REPORT DATE: 11 Nov[redacted]76

13. NUMBER OF PAGES: 28 pages

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 12 32p.

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report):

Approved for public release; distribution unlimited

14 DGS-249, Scientific-6

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):

16 7600 17 03

18. SUPPLEMENTARY NOTES:

TECH, OTHER

19. KEY WORDS (Continue on reverse side if necessary and identify by block number):

geodesy; gravitational field; inverse problems, geophysical; least-squares methods

20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

The report compares least-squares collocation methods for determining the earth's gravitational field with geophysical inversion techniques. Both are underdetermined problems with strong structural similarities. Collocation is also considered from the point of view of representing the external gravitational field by means of analytic functions. Finally, the conceptual basis of least-squares collocation is briefly discussed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

400254

HB

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

FOREWORD

This report was prepared by Dr. Helmut Moritz, Professor Technical University at Graz and Adjunct Professor, Department of Geodetic Science of The Ohio State University, under Air Force Contract No. F19628-76-C-0010, The Ohio State University Research Foundation, Project No. 4214A1, Project Supervisor, Urho A. Uotila, Professor, Department of Geodetic Science. The contract covering this research is administered by the Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts, with Mr. Bela Szabo, Contract Monitor.

CONTENTS

1. Introduction

In geodesy and geophysics we frequently meet with the situation that a model defined by a set of, say, N parameters is to be determined from a smaller number $n < N$ of observations.

As an example, the internal structure of the earth may be defined by a set of N parameters describing the density, the rigidity, and the compressibility of the earth as a function of depth. The n observations comprise velocities of seismic surface waves, together with the mass and the polar moment of inertia of the earth. If the model for the earth's internal structure is to be realistic, then N will be large and $n < N$.

We thus have $n < N$ equations for N unknowns, which is obviously an underdetermined problem admitting an infinite number of possible solutions. Using standard mathematical terminology (Lanczos, 1964), we have an improperly posed problem. (A problem is properly posed if it has a unique solution that depends continuously on the data.)

Originally, the equations expressing the data x_i as functions of the model parameters s_r will, in general, be non-linear:

$$x_i = f_i(s_1, s_2, \ldots, s_N) , \qquad i = 1,2,\ldots,n . \tag{1-1}$$

By a suitable application of Taylor's theorem it is usually possible to approximate these equations by linear ones:

$$x_i = \sum_{r=1}^{N} a_{ir} s_r , \tag{1-2}$$

or in matrix notation:

$$\underline{x} = \underline{As} . \tag{1-3}$$

The formal solution of this system of linear equations may be written as

$$\underline{s} = \underline{A}^{-1}\underline{x} \; . \tag{1-4}$$

If $\underline{A}$ were a regular square matrix, then $\underline{A}^{-1}$ would be the ordinary inverse matrix of $\underline{A}$. In our underdetermined case, however, $\underline{A}^{-1}$ must be understood in the sense of generalized matrix inverses (cf. Bjerhammar, 1973; Rao and Mitra, 1971).

At any rate, the solution of (1-1) or (1-3) may be considered as an inversion of these equations with respect to the parameters s_r, which accounts for the name, geophysical inverse problems.

Another typical example of an "improperly posed" inverse problem is the determination of subsurface mass distributions which produce a given anomalous gravity field at the earth's surface. This problem is sometimes called an inverse problem of potential theory (Lavrentiev, 1967; Burkhard and Jackson, 1976).

The determination of the earth's external gravitational field from geodetic, gravimetric and satellite data may also be considered as an inverse problem that is mathematically quite similar to the determination of the internal structure of the earth from seismic and other data.

This geodetic inverse problem is likewise underdetermined. The external gravitational field requires for a complete description an infinite number of parameters, for instance, the set of all coefficients in the expansion of the external gravitational potential in spherical harmonics. This infinite number, $N = \infty$, of parameters is to be determined from a finite number n of observations.

Even in the seismic inverse problem it is, at least theoretically, appropriate to take $N = \infty$ if we wish to admit reasonably general functions for density, rigidity, and compressibility because it cannot be assumed *a priori* that such functions depend on a finite number of parameters only.

Thus, in general, the space of parameters will be infinite-dimensional rather than N-dimensional. In other words, the proper general setting for (linear) geodetic and geophysical inverse problems will be infinite-dimensional Hilbert space. This was pointed out by Krarup (1969) for the geodetic case and by Backus (1970) for geophysical inverse problems.

The geodetic inverse problem, the determination of the external gravitational field from data of different kind, is usually solved by least-squares collocation. This technique has many features in common with other geophysical inversion methods. It may, therefore, be of interest to compare these techniques and to exhibit some cross-connections.

We shall also discuss least-squares collocation from the point of view of analytically representing the external gravitational field by a linear combination of suitable simpler harmonic functions.

The subject of the present report is purely conceptual, aiming at a better understanding of least-squares collocation by considering it in is relation to other methods; no new computational formulas will be derived. Still, this paper might be useful as a contribution to the present discussion on the conceptual foundations of least-squares collocation.

2. Systems of Linear Equations

Let us assume that the geophysical or geodetic inversion problem has already been linearized, so that it reduces to the solution of a system of linear equations of form (1-3),

$$\underline{B}\underline{s} = \underline{x} \, , \tag{2-1}$$

$\underline{s}$ being the vector of N parameters to be determined, $\underline{x}$ denoting the vector of n observations, and $\underline{B}$ being a given $n \times N$ matrix of coefficients.

Assuming $n < N$, we have an underdetermined problem. The number N of parameters may be finite or infinite. For $N = \infty$ we must presuppose that the occurring sums from 1 to N, which are now infinite series,converge; otherwise the formal operations are the same as for a finite N.

Equations of type (2-1) may be formulated for seismic inversion problems (Knopoff and Jackson, 1973), for gravity interpretation problems (Burkhard and Jackson, 1976; Kaula et al., 1975), as well as for the geodetic problem of determining spherical harmonics of the geopotential from satellite observations by least-squares collocation (Schwarz, 1974, 1975).

A general solution of the *underdetermined system* (2-1) (assumed full rank) is

$$\underline{s} = \underline{C}\underline{B}^T(\underline{B}\underline{C}\underline{B}^T)^{-1}\underline{x} \, , \tag{2-2}$$

where $\underline{C}$ is a $N \times N$ matrix such that $\underline{BCB}^T$ is a regular $n \times n$ matrix. Otherwise $\underline{C}$ is arbitraty: different solutions are obtained by different choices of $\underline{C}$, and this gives an infinite set of possible solution vectors $\underline{s}$.

It is immediately verified that (2-2) satisfies the given system (2-1). Less obvious but also well known from the theory of generalized inverses (cf. Bjerhammar, 1973, p.116) is the fact that the solution (2-2) satisfies the minimum condition

$$\underline{s}^T\underline{C}^{-1}\underline{s} = \text{minimum} \, , \tag{2-3}$$

provided the inverse matrix $\underline{C}^{-1}$ exists in an appropriate sense (for $N = \infty$ this implies convergence of any occurring infinite sums).

Usually the measurements x_i will be affected by unknown observational errors, denoted by n_i ; the notation follows the terminology of time series: "$\underline{s}$" stands for signal, and "$\underline{n}$" stands for "noise". (There is hardly any danger of confusing $\underline{n}$, the noise vector, with n , the number of observations.) Then (2-1) is to be replaced by

$$\underline{Bs} + \underline{n} = \underline{x} . \tag{2-4}$$

An appropriate minimum condition, instead of (2-3), is now

$$\underline{s}^T\underline{C}^{-1}\underline{s} + \underline{n}^T\underline{D}^{-1}\underline{n} = \text{minimum}, \tag{2-5}$$

where $\underline{C}$ and $\underline{D}$ are symmetrical matrices that can be interpreted statistically as <u>covariance matrices</u>: $\underline{C}$ is the covariance matrix of the signal $\underline{s}$, and $\underline{D}$ is the covariance matrix of the noise $\underline{n}$, that is, of the observational errors.

The solution of (2-4) under the minimum condition is found to be

$$\underline{s} = \underline{CB}^T(\underline{BCB}^T + \underline{D})^{-1}\underline{x} . \tag{2-6}$$

This solution has been used for geophysical inverse problems (Wiggins, 1972, pp.260-1; Burkhard and Jackson, 1976, p.1514).

It is also the solution obtained by least-squares collocation (Schwarz, 1974). This is clear from the condition (2-5) and may also be derived directly as follows.

The basic collocation model is

$$\underline{x} = \underline{AX} + \underline{s}' + \underline{n} \tag{2-7}$$

(Moritz, 1972, p.7), where $\underline{s}'$ is the random "signal part" of the observations $\underline{x}$ and $\underline{X}$ denotes non-random (systematic) parameters. If $\underline{X} = 0$, the model (2-7) reduces to

$$\underline{x} = \underline{s}' + \underline{n} . \tag{2-8}$$

If the vector $\underline{s}'$ is expressed as a linear combination $\underline{Bs}$ of "basic signals" (in the geodetic case, e.g., spherical harmonic coefficients), then (2-8) becomes

$$\underline{x} = \underline{Bs} + \underline{n} , \tag{2-9}$$

which, in fact, is the model (2-4), the "parameters" $\underline{s}$ now being treated as random variables.

The collocation solution follows from eq. (2-36) of (Moritz, 1972, p.15) for $\underline{X} = 0$:

$$\underline{s} = \underline{C}_{sx}\underline{C}_{xx}^{-1}\underline{x} , \tag{2-10}$$

which is essentially nothing but the fundamental Wiener-Kolmogorov prediction formula (cf. Liebelt, 1967; he calls it Gauss-Markov theorem). The matrix $\underline{C}_{xx}$ is the covariance matrix of the vector $\underline{x}$, and $\underline{C}_{sx}$ is the cross-covariance matrix between the vectors $\underline{s}$ and $\underline{x}$.

We assume the vectors $\underline{s}$ and $\underline{n}$ to be uncorrelated. Then the application of covariance propagation (Moritz, 1972, p.97) gives readily.

$$\underline{C}_{xx} = \underline{BCB}^T + \underline{D} , \tag{2-11}$$

$$\underline{C}_{sx} = \underline{CB}^T , \tag{2-12}$$

$\underline{C}$ being the covariance matrix of the vector $\underline{s}$. Thus, in our case, (2-10) in fact becomes (2-6).

Let us now turn to the accuracy of the estimated signal $\underline{s}$. This accuracy is usually defined by the error covariance matrix $\underline{E}_{ss}$ consisting of error variances (squares of standard errors) as diagonal terms and error covariances as off-diagonal terms.

If we abbreviate (2-6) as

$$\underline{s} = \underline{Lx} \, , \tag{2-13}$$

then $\underline{E}_{ss}$ is given by eq. (3-20) of (Moritz, 1972, p.29):

$$\underline{E}_{ss} = \underline{C}_{ss} - \underline{LC}^T_{sx} - \underline{C}_{sx}\underline{L}^T + \underline{LC}_{xx}\underline{L}^T \, . \tag{2-14}$$

With $\underline{C}_{ss} = \underline{C}$ and (2-11) and (2-12), this becomes

$$\underline{E}_{ss} = \underline{C} - \underline{LBC} - \underline{CB}^T\underline{L}^T + \underline{L}(\underline{BCB}^T + \underline{D})\underline{L}^T \, .$$

This is readily given the form

$$\underline{E}_{ss} = \underline{E}_1 + \underline{E}_2 \, , \tag{2-15}$$

where

$$\underline{E}_1 = (\underline{LB} - \underline{I})\underline{C}(\underline{LB} - \underline{I})^T \, , \tag{2-16}$$

$$\underline{E}_2 = \underline{LDL}^T \, , \tag{2-17}$$

$\underline{I}$ denoting the unit matrix. This form provides an elegant decomposition of the error covariance matrix $\underline{E}_{ss}$ which has the following interpretation (Burkhard and Jackson, 1976, p.1514).

Let us write (2-13) in the form

$$\underline{\hat{s}} = \underline{Lx} \ , \tag{2-18}$$

where the notation $\underline{\hat{s}}$ is to indicate that we are dealing with the estimated value of the signal. We now substitute into this formula eq. (2-4),

$$\underline{x} = \underline{Bs} + \underline{n} \ , \tag{2-19}$$

in which $\underline{s}$ and $\underline{n}$ denote the "true" values. The result is

$$\underline{\hat{s}} = \underline{LBs} + \underline{Ln} \ , \tag{2-20}$$

so that the "true error" of the signal, defined as estimated value $\hat{s}$ minus true value s, is given by

$$\underline{\hat{s}} - \underline{s} = (\underline{LB} - \underline{I})\underline{s} + \underline{Ln} \ . \tag{2-21}$$

The first term on the right-hand side,

$$\underline{e}_1 = (\underline{LB} - \underline{I})\underline{s} \ , \tag{2-22}$$

denotes the resolving error, due to the deviation of the product $\underline{LB}$ from the unit matrix $\underline{I}$: if the system (2-1) could be exactly solved--$\underline{B}$ being a regular square matrix--then $\underline{e}_1$ would be zero. The second term,

$$\underline{e}_2 = \underline{Ln} \ , \tag{2-23}$$

expresses simply the effect of data errors propagating into the solution. Clearly, $\underline{E}_1$ is the covariance matrix of $\underline{e}_1$ and $\underline{E}_2$ the covariance matrix of $\underline{e}_2$; $\underline{E}_1$ and $\underline{E}_2$ simply add in (2-15) because $\underline{e}_1$ and $\underline{e}_2$ are uncorrelated, being linear functions of the uncorrelated vectors $\underline{s}$ and $\underline{n}$, respectively.

We have at length considered underdetermined systems, because they are typical for geophysical inversion problems and for the determination of the gravitational field. As regards (full-rank) overdetermined systems, the model (2-4) can still be used with $\underline{B}$ as a "standing" instead of a "lying" rectangular matrix, the vector $\underline{s}$ having now less components than the vector $\underline{x}$, $N < n$. The condition (2-5) gives again a solution which is formally identical to (2-6).

The new feature, peculiar to the overdetermined case $n > N$, is that now the following well-known transformation (cf. Liebelt, 1967, p.30) can be applied:

$$\underline{C}\underline{B}^T(\underline{B}\underline{C}\underline{B}^T+\underline{D})^{-1} = (\underline{B}^T\underline{D}^{-1}\underline{B}+\underline{C}^{-1})^{-1}\underline{B}^T\underline{D}^{-1} . \tag{2-24}$$

Thus, for the overdetermined case, (2-6) is equivalent to

$$\underline{s} = (\underline{B}^T\underline{D}^{-1}\underline{B}+\underline{C}^{-1})^{-1}\underline{B}^T\underline{D}^{-1}\underline{x} . \tag{2-25}$$

This form has the advantage of permitting an easy transition to least-squares adjustment by parameters: the condition for this case,

$$\underline{n}^T\underline{D}^{-1}\underline{n} = \text{minimum}, \tag{2-26}$$

arises from (2-5) by letting $\underline{C}^{-1} \to 0$; in the same way, (2-25) becomes

$$\underline{s} = (\underline{B}^T\underline{D}^{-1}\underline{B})^{-1}\underline{B}^T\underline{D}^{-1}\underline{x} , \qquad (2\text{-}27)$$

which is the classical least-squares solution of (2-4) if it is an overdetermined system.

It may be mentioned that the two elementary cases of classical least-squares adjustment, namely adjustment by parameters and adjustment by conditions, come out as limiting cases of the general model (2-6). They are, so to speak, at extreme ends of the scale: for overdetermined systems, (2-6) gives (2-27) for $\underline{C}^{-1} \to 0$, as we have just seen. For underdetermined systems, (2-6) becomes (2-2) on letting $\underline{D} \to 0$. However, (2-6) is formally identical to the result of least-squares adjustment by conditions, on interpreting $\underline{s}$ now as measuring errors; cf. eq. (12.1:11) of(Bjerhammar, 1973, p.166) with our eq. (2-2) for $\underline{C} = \underline{I}$.

Complete Sets of Solutions.- An interesting question is the problem of determining the set of all "reasonable" gravitational fields that are compatible with the set of all available geodetic data. Each measurement (such as terrestrial gravity or satellite data) gives an equation, which, on linearization, provides one of the linear equations of a system (2-1); we assume n independent observations. The unknowns $\underline{s}$ are to be the set of all zonal and tesseral harmonics in the anomalous gravitational field; thus $N = \infty$.

For simplicity we assume errorless data $(\underline{n} = 0)$; thus the equations (2-1) have to be satisfied exactly. As we have mentioned, such a complete set of solutions is obtained by letting the matrix $\underline{C}$ vary in the solution (2-2). However, this representation is computationally inconvenient since each time a large matrix $\underline{BCB}^T$ would have to be inverted anew.

Practically more suitable is the representation, well-known from the theory of singular matrix inverses (cf. Bjerhammar, 1973, p.378),

$$s = B^- x + (I - B^- B) u \ , \tag{2-28}$$

where

$$\underline{B}^- = \underline{C}_o \underline{B}^T (\underline{B}\underline{C}_o \underline{B}^T)^{-1} \tag{2-29}$$

is a fixed generalized inverse of $\underline{B}$, to be computed once for all with an assumed fixed matrix $\underline{C}_o$. The vector $\underline{u}$ is arbitrary; the desired set of solutions is obtained by letting $\underline{u}$ run through the set of all possible vectors of infinitely many components for which the infinite sums converge.

Having thus obtained the complete set of solutions (2-28), one may select special solutions by imposing suitable conditions. The geodetically most meaningful condition is probably optimal accuracy of the result, expressed in terms of minimum error variances. This condition is known to be equivalent, for errorless data, to the condition (2-3), $\underline{C}$ being the signal covariance matrix (cf. Moritz, 1972, pp.38 and 122). Under this condition the solution is given by (2-2). For this case we may put

$$\underline{u} = \left[\underline{C}\underline{B}^T (\underline{B}\underline{C}\underline{B}^T)^{-1} - \underline{B}^- \right] \underline{x} \ ; \tag{2-30}$$

in fact, on substituting this expression, (2-28) becomes (2-2).

A geophysically interesting side condition is that the solution $\underline{s}$ belong to a <u>convex set</u> described by inequalities such as

$$s_1 \geqq 0 \ , \quad 2s_1 + s_2 \geqq 5 \ , \text{ etc.}$$

Problems of minimizing the error variance (or other functions) subject to such side conditions are dealt with in optimization theory (linear and nonlinear programming).

In the problem of determining mass structures from gravity data, the requirement that the obtained density be nonnegative would lead to such side conditions.

In the geodetic inversion problem, the determination of the external gravitational field, such conditions have never been used and do not seem, in general, to have practical importance. However, the condition that the total gravitational field should be generated by masses that are everywhere nonnegative, is a physical requirement to be met. In "reasonable" solutions it seems to be more or less automatically satisfied, but a closer look on the convexity problem in geodetical applications may not be without theoretical interest, as has been pointed out to the author recently by Prof. P.C. Sabatier and Prof. D.D. Jackson.

We finally mention the beautiful geometrical treatment of underdetermined and overdetermined systems of linear equations in (Lanczos, 1964, chapter 3).

3. Functional Representation of the Gravitational Field

Let us now turn to an apparently quite different problem, the representation of the earth's external gravitational field by appropriate analytical functions.

Let the anomalous potential T be approximated by a linear combination f of suitable base functions $\phi_1, \phi_2, \phi_3, \ldots$:

$$T(P) = f(P) = \sum_{k=1}^{n} b_k \phi_k(P) , \tag{3-1}$$

P denoting the space point at which these functions are being considered, and b_k denoting suitable coefficients.

Since T is harmonic outside the earth's surface, all base functions ϕ_k must be harmonic functions, too. They may, for instance, be spherical harmonics or the potentials of point masses suitably distributed below the earth's surface.

How are the coefficients b_k to be determined? Assume, for the sake of simplicity and definiteness, that we are given errorless values of T at n space points P_i. Then it is reasonable to use the approximation of T by a linear combination (3-1) of n base functions and to postulate that f exactly reproduces T at the n given points. Putting

$$T(P_i) = f(P_i) = f_i \ , \tag{3-2}$$

we thus have the conditions

$$\sum_{k=1}^{n} b_k \phi_k(P_i) = f_i \ , \tag{3-3}$$

which are n linear equations for the n unknowns b_k, which can be solved provided they are linearly independent.

With

$$\phi_k(P_i) = A_{ik} \tag{3-4}$$

we thus have

$$\sum_{k=1}^{n} A_{ik} b_k = f_i \tag{3-5}$$

or in matrix notation,

$$\underline{Ab} = \underline{f} \ , \tag{3-6}$$

with the solution

$$\underline{b} = \underline{A}^{-1}\underline{f} . \tag{3-7}$$

Thus the determination of the coefficients b_k always involves the solution of a $n \times n$ system of linear equations, or the inversion of a $n \times n$ matrix.

An exception would be the case that the matrix $\underline{A}$ reduces to the unit matrix, that is

$$\chi_k(P_i) = \delta_{ik} = \begin{cases} 1 , & i = k \\ 0 , & i \neq k \end{cases} , \tag{3-8}$$

the base functions for this case being denoted by $\chi(P)$ instead of $\phi(P)$. This means that the functions χ_k are zero at all data points except one, at which they assume the value 1 . Such functions are called <u>sample functions</u>; they have the pleasant property that, by (3-7), $\underline{b} = \underline{f}$, so that (3-1) reduces to

$$f(P) = \sum_{k=1}^{n} f_k \chi_k(P) , \tag{3-9}$$

the coefficients being simply the given values of the function f at the data points P_k .

Of course, it cannot be expected in general that $\underline{A}$ in fact reduces to the unit matrix, but even if this is not the case, we can associate a set of sample functions χ_k to the given base functions ϕ_k , by putting

$$\chi_k(P) = \sum_{j=1}^{n} A_{jk}^{(-1)} \phi_j(P) , \tag{3-10}$$

where $A_{jk}^{(-1)}$ denote the elements of the matrix $\underline{A}^{-1}$ inverse to $\underline{A}$. In fact, putting $P = P_i$ in (3-10) we get

$$\chi_k(P_i) = \sum_{j=1}^{n} A_{jk}^{(-1)} A_{ij}$$

$$= \sum_{j=1}^{n} A_{ij} A_{jk}^{(-1)} = \delta_{ik} , \tag{3-11}$$

according to the definition of the inverse matrix, so that (3-8) is satisfied. Furthermore, the substitution of (3-10) into (3-9) gives

$$f(P) = \sum_{j=1}^{n} \sum_{k=1}^{n} A_{jk}^{(-1)} f_k \phi_j(P)$$

$$= \sum_{j=1}^{n} b_j \phi_j(P) \tag{3-12}$$

by (3-7), so that the sample function development (3-9) is identical to the original representation (3-1).

In this way we can, indeed, associate to each system of base functions $\phi_1, \phi_2, \ldots$ and to each configuration of points $P_1, P_2, \ldots, P_n$ a system of sample functions $\chi_1, \chi_2, \ldots, \chi_n$. This has advantages if the same point configuration is used with different data sets; then only the f_k change, but the sample functions $\chi_k(P)$ remain the same. As a matter of fact, also this approach requires, according to (3-10), the inversion of the $n \times n$ matrix $\underline{A}$, but it need to be performed only once for a given point configuration.

The type of sample functions best known in geodesy are the functions of Giacaglia and Lundquist (1972) based on the system of spherical harmonics for regular point configurations on the sphere. For such point configurations it presents computational advantages and is precisely equivalent to a spherical-harmonic representation up to a certain degree.

Spherical harmonics and sample functions based on them are suitable for a global representation. For local representations, potentials of buried masses or related functions (Dufour and Kovalevsky, 1970) are more suitable; another possibility is least-squares interpolation to be considered now.

The approximation of the potential T by a finite linear combination of n base functions ϕ_k can, of course, reproduce T only at n points, the data points. At other points, f will deviate from T, that is, the error

$$\varepsilon_P = T(P) - f(P) \tag{3-13}$$

will, in general, differ from zero. Of considerable importance is obviously the question whether there exists a system of base functions $\phi_k(P)$ for which, at every point P, the mean square interpolation error m_P, defined by

$$m_P^2 = M(\varepsilon_P^2) \tag{3-14}$$

attains a minimum, M denoting a suitably defined statistical expectation value.

The answer to this question is yes (which is by no means a matter of fact). This leads to least-squares interpolation due to N. Wiener and A.N. Kolmogorov. A derivation is found in (Heiskanen and Moritz, 1967, p.268). The result is

$$\phi_k(P) = C(P, P_k) , \tag{3-15}$$

where $C(P,P_k)$ is the covariance function of T between points

P and P_k. Now this covariance function admits the representation (Moritz, 1972, p.88)

$$C(P,Q) = \sum_{n=0}^{\infty} k_n \left(\frac{R^2}{r_P r_Q}\right)^{n+1} P_n(\cos\psi) , \tag{3-16}$$

where k_n and R are constants, r_P and r_Q are the spatial radius vectors of P and Q, and P_n are Legendre's polynomials as functions of the angle ψ between r_P and r_Q. This representation shows that $C(P,P_k)$, considered as a function of P, is a harmonic function, analytic everywhere outside a certain sphere. If the radius of this sphere is chosen so that the sphere lies completely inside the earth, then the functions (3-15) will be analytic everywhere outside and on the earth's surface and can, therefore, be used for representing the external gravitational potential.

It is a general principle of least-squares estimation that the result does not depend strongly on the a priori covariances used. Therefore, if the actual covariance function (3-16) is a function which is too complicated for practical use, one may, without appreciable loss of accuracy, replace the actual function by a suitable approximation of a simpler and more manageable analytical form (in practice, one will anyway fit a rather simple analytical expression to empirically obtained covariance data). Therefore, also the functions (3-15) will be precisely defined harmonic functions of a relatively simple analytical form.

By (3-4) and (3-15) we have now

$$A_{ik} = C(P_i,P_k) \stackrel{\text{def}}{=} C_{ik} . \tag{3-17}$$

Defining the matrix $\underline{C}$ by

$$\underline{C} = (C_{ik}) = \underline{A} , \tag{3-18}$$

we get from (3-7)

$$\underline{b} = \underline{C}^{-1}\underline{f} \;. \tag{3-19}$$

On introducing the vector

$$\underline{C}(P) = \left[\, C(P,P_1)\ ,\ C(P,P_2)\ ,\ldots,\ C(P,P_n) \,\right] \;, \tag{3-20}$$

(3-1) thus takes the form

$$f(P) = \underline{C}(P)\underline{C}^{-1}\underline{f} \;, \tag{3-21}$$

which is again the Wiener-Kolmogorov prediction formula (2-10).

It is clear that, on the basis of the functions (3-15) for least-squares prediction, *corresponding sample functions* can be defined by (3-10), each of which is zero at all data points except one, at which it takes the value 1.

Least-squares prediction shares with all representation methods of type (3-1) the disadvantage that a (usually large) $n \times n$ matrix has to be inverted; it has, however, the advantage that in this particular case the matrix to be inverted is symmetric, being a covariance matrix (in the general case, the matrix $\underline{A}$ is not symmetric).

Taken in the form just described, the application of least-squares prediction in geodesy is rather limited, essentially to the prediction of gravity anomalies. In fact, however, geodetic data are of many different types: horizontal and vertical angles, distances, gravity measurements, deflections of the vertical, satellite observations of many kinds, gravity gradients, etc. All these geodetic data obviously share the properties that they depend on the earth's gravitational field; therefore, on linearization, all of them may be expressed a linear functionals of the anomalous potential T. (Linear Functionals are by no means restricted to linear functions but they include differential and integral formulas such as Stokes' integral.)

So more general and geodetically more important is the problem to fit the representation (3-1) to the data, so that the n given functionals of T are exactly reproduced. This is the principle of collocation, which is frequently used in numerical mathematics (cf. Collatz, 1966). If the base functions $\phi_k(P)$ are again to be determined by the least-squares condition of minimum interpolation error at all points P, then it is natural to call the method least-squares collocation.

The solution has again the same formal structure as in least-squares prediction. The result is

$$\phi_k(P) = C(P, x_k) , \tag{3-22}$$

which is the covariance between $T(P)$ and the measurement x_k; it is a function of the point P that is again harmonic and analytic. The coefficients b_k forming the vector $\underline{b}$ are determined by

$$\underline{b} = \underline{C}_{xx}^{-1} \underline{x} , \tag{3-23}$$

where $\underline{C}_{xx}$ is the autocovariance matrix of the observation vector $\underline{x} = (x_k)$.

Thus, with errorless data, least-squares collocation determines the analytical form of the functions ϕ_k by the requirement of optimal accuracy, whereas the data are exactly reproduced.

In general, the geodetic data will be affected by random measuring errors. Also in this case the condition m_P = minimum can be applied and determines simultaneously

(1) the best analytical expression for the functions ϕ_k and

(2) the best values for the coefficients b_k .

Even in this case, the formal expressions are the same as before, again being given by (3-22) and (3-23). The measuring errors have no influence on the choice of ϕ_k by (3-22), so that ϕ_k again represent pure gravitational covariance functions, that is, analytical and harmonic functions; again the least-squares principle serves only to single out the most suitable analytical expression for the base functions ϕ_k among the many possible choices.

Where statistics enters directly is the determination of the coefficients b_k , which is done in such a way that the effect of random data errors is minimized (therefore, the data are no longer exactly reproduced); in statistical terminology, we have a "best linear estimate": an *unbiased* estimate of minimum variance.

Expressions analogous to (3-1) may be given for any other quantity of the anomalous gravitational field (called "signal"), such as geoidal heights, deflections of the vertical, gravity anomalies, etc. The coefficients b_k remain the same since they depend only on the data $\underline{x}$ by (3-23). What changes are the base functions ϕ_k ; the new base functions are derived by simple analytical operations such as differentiation, since a linear operation performed on (3-1) acts on the base functions ϕ_k only.

Since these base functions are covariance functions, these operations are special cases of covariance propagation which plays a basic role in least-squares collocation: it has to carry the precise mathematical structure of the gravitational field (cf. Moritz, 1972, pp.94-99).

<u>Inclusion of Systematic Parameters</u>.- A last restriction has to be removed before least-squares collocation can be fully applied to general geodetic problems. So far we have assumed that

we deal only with quantities that have zero statistical expectation such as the elements of the anomalous gravitational field, so that systematic effects had to be removed beforehand; we shall now free ourselves from this restriction.

We shall, therefore, consider the following model:

$$\underline{x} = \underline{AX} + \underline{s} + \underline{n} \,, \tag{3-24}$$

where the vector $\underline{x}$ comprises the measured quantities, $\underline{s}$ representing the part due the anomalous gravitational field and $\underline{n}$ denoting random measuring errors. The new component is $\underline{AX}$, where the vector $\underline{X}$ comprises systematic, nonrandom parameters and $\underline{A}$ is a given matrix of coefficients.

This model is general enough to encompass all conceivable geodetic measurements. In fact, any geodetic measurement may be split up, according to (3-24), into three parts:

1. A systematic part $\underline{AX}$ involving, on the one hand, the ellipsoidal reference system and, on the other hand, other parameters and systematic errors (originally non-linear functions are again thought to have been linearized by Taylor's theorem);
2. A random part $\underline{s}$ (of zero expectation) expressing the effect of the anomalous gravity field; and
3. Random measuring errors $\underline{n}$.

As an example, consider a measurement of gravity, g . Here $\underline{AX}$ represents normal gravity γ , as well as systematic errors such as gravimeter drift; $\underline{s}$ is the gravity anomaly Δg ; and $\underline{n}$ stands for the measuring error. Other examples will be found in (Moritz, 1972, pp.70-76).

The formulas for estimating $\underline{X}$ and $\underline{s}$ may be derived from two different but equivalent minimum principles:

1. From a least-squares principle corresponding to (2-5);
2. From the condition of minimum variance (least standard

errors for estimated $\underline{X}$ and $\underline{s}$).

The result is:

$$\underline{X} = (\underline{A}^T \underline{C}_{xx}^{-1} \underline{A})^{-1} \underline{A}^T \underline{C}_{xx}^{-1} \underline{x} \ , \qquad (3\text{-}25)$$

$$\underline{s} = \underline{C}_{sx} C_{xx}^{-1} (\underline{x} - \underline{AX}) \ . \qquad (3\text{-}26)$$

The first equation is analogous to classical least-squares adjustment by parameters, except that the covariance matrix $\underline{C}_{xx}$ includes now covariances of the signal as well as those of the measuring errors. The second equation is a fairly obvious generalization of the Kolmogorov-Wiener formula (2-10) or (3-21) to the case in which the expectation of $\underline{x}$ is $\underline{AX}$ rather than zero.

These formulas are an extension of the corresponding results for time series (Grenander and Rosenblatt, 1957, p.87).

The present method, least-squares collocation with parameters, may be considered as a combination of least squares adjustment and least-squares prediction into a unified scheme, which makes possible the use of all conceivable geodetic data--classical angle and distance measurements, gravity measurements, satellite data of different kind, etc.--to obtain the geometric position of points of the earth's surface as well as the gravitational field.

4. The Many Facets of Collocation

The Nature of Least-Squares Collocation.- In sec.2 we have seen that, in important cases, least-squares collocation reduces to a least-squares solution of an underdetermined system of linear equations. It might thus seem that least-squares collocation is essentially nothing else than classical least-squares methods known from adjustment computations.

In trying to answer this question, we shall first consider "simple" least-squares collocation, that is without estimation of systematic parameters, so that we have $\underline{A} = 0$ in (3-24) and (3-26); this was also done in sec.2. Physically this means that the observations $\underline{x}$ are all quantities of the anomalous gravity field.

In fact, every problem of simple least-squares collocation can, in principle, be formulated as a system of linear equations of the form discussed in sec.2. Every quantity of the anomalous gravity field, for instance a geoidal height, a gravity anomaly, or a deflection of the vertical, can be expressed as an infinite series of spherical harmonics, the coefficients of which constitute the unknowns, the number N of which is obviously infinite. (Convergence problems may be overcome by the use of Runge's theorem (Krarup, 1969, p.54; Moritz, 1971, p.79).) Thus we always get a system of equations of type (2-9), as outlined in sec.2.

If N were finite, then the solution (2-6) would, in fact, correspond to classical least-squares since the condition (2-5) can be written in the form

$$\underline{v}^T \underline{P} \underline{v} = \text{minimum} \tag{4-1}$$

with

$$\underline{v} = \begin{bmatrix} \underline{s} \\ \\ \underline{n} \end{bmatrix}, \qquad \underline{P} = \begin{bmatrix} \underline{C}^{-1} & 0 \\ \\ 0 & \underline{D}^{-1} \end{bmatrix} \tag{4-2}$$

and (2-4) takes the form of a condition equation for the vector $\underline{v}$.

What difference does the fact $N = \infty$ make? The signal space is now no longer finite-dimensional but it is infinitely-dimensional Hilbert space. The elements of Hilbert spaces may be infinite vectors, such as $\underline{s}$, but also functions of a certain kind. In the gravitational case, we conveniently consider spaces of harmonic functions, more precisely, of functions harmonic outside a certain sphere.

There is a one-to-one correspondence between such functions and infinite vectors $\underline{s}$ (of norm $\underline{s}^T\underline{s} < \infty$) by taking $\underline{s}$ as the vector of coefficients of the spherical-harmonic expansion. This makes it possible to avoid the use of infinite vectors (and corresponding convergence problems) by working with functions and linear operations on them. In this way, formulas of Kolmogorov-Wiener type (3-21) are obtained, which operate with finite n x n matrices.

Another way of linking collocation (including systematic parameters) with adjustment, this time in finite-dimensional space, was given in (Moritz, 1972, pp. 12-15). There, however, the signal values to be computed do not enter in the condition equations although they do enter into the minimum principle $\underline{v}^T\underline{Pv}$. Thus the condition equations do not provide a complete formulation as they do in adjustment computations, but must be supplemented by additional considerations: the new signals are related to the observations not by the condition equations, but through joint covariances. (The importance of joint covariances is well known from wide-sense stationary stochastic processes where they carry the total statistical structure.)

So the relation of collocation and adjustment is very elusive: the analogies are striking, but at the very moment when we think that we have hit on an exact identity we must recognize a difference in a fine but essential point. For a more detailed discussion see also (Rummel, 1976).

A much closer relationship is between least-squares collocation and the theory of prediction of stationary stochastic processes. In fact, the anomalous gravity potential may be considered as a stochastic process on the sphere, or rather as a spatial stochastic process that is harmonic outside a sphere.

The theory of stochastic processes provides a very convenient mathematical formalism and a statistical interpretation and terminology (e.g., covariance functions). In fact, we have seen at the end of the preceding section that the basic collocation formulas (3-25) and (3-26) are precise analogues of the corresponding results for prediction and parameter estimation of time series.

Some kind of statistical interprediction is desirable to get the geodetically important concept of accuracy (standard errors of prediction) into the picture. How seriously the statistical interpretation is taken, is a matter of controversy and also of personal taste. The present author (Moritz, 1972, secs.8 and 9) favors an interpretation in terms of Norbert Wieners "covariance of individual functions" to take into account the fact that there is only <u>one</u> earth and to avoid difficulties associated with the stochastic process interpretation as pointed out by Lauritzen (1973): it is impossible to find a stochastic process, harmonic outside the sphere, that is both Gaussian and ergodic.

It is possible largely to play down the statistical aspects, emphasizing Hilbert space geometry and considering the covariance function as a kernel function in Hilbert space, as Krarup did in his fundamental paper (1969). In fact, an interpolation formula formally identical to the Kolmogorov-Wiener prediction formula is obtained in the theory of kernel functions in Hilbert space (Meschkowski, 1962, p.114) in a completely "deterministic" way, without using any statistics.

It is well to emphasize the analytical nature of kernel functions--which are precisely defined harmonic functions--to avoid the wrong impression that least-squares collocation "messes up everything statistically" and to show that we have an underlying completely "clean" analytical model. The elementary reasoning in sec.3 of the present paper is intended to serve the same purpose.

Why, then, use a special name, "least-squares collocation" for the present geodetic method and not simply call it prediction? The term prediction is usually understood to comprise interpolation and extrapolation of time series and data of the same kind, for instance, gravity anomalies. The essence of the present method, however, is the use of heterogeneous data, which are linear functionals of the anomalous gravitational potential. As we have seen in the preceding section, the fitting of a function to given functionals is precisely the feature of collocation as understood in numerical mathematics.

The name "least-squares collocation" thus seems to be quite suitable to characterize the present method as a representation of the anomalous gravitational field by "clean" analytical functions, which are selected and the coefficients of which are determined by a least-squares principle.

REFERENCES

Backus, G. (1970) Inference from inadequate and inaccurate data, I and II. Proc. Natl. Acad. Sci., 65, 1-7 and 281-287.

Backus, G., and Gilbert, F. (1968) The resolving power of gross earth data. Geophys. J.R. Astr. Soc., 16, 169-205.

Bjerhammar, A. (1973) Theory of Errors and Generalized Matrix Inverses. Elsevier, Amsterdam.

Burkhard, N., and Jackson, D.D. (1976) Application of stabilized linear inverse theory to gravity data. J. Geophys. Res., 81, 1513-1518.

Collatz, L. (1966) The Numerical Treatment of Differential Equations. Springer, Berlin.

Dufour, H.M., and Kovalevsky, J. (1970) Formulation pratique du champ de gravitê terrestre par des fonctions rêgionalisêes. Institut Gêographique National Paris, Report No. 26.837.

Giacaglia, G.E.O., and Lundquist, C.A. (1972) Sampling functions for geophysics. Smithsonian Astrophys. Observatory, Spec. Report No. 344.

Grenander, V., and Rosenblatt, M. (1957) Statistical Analysis of Stationary Time Series. Wiley, New York.

Heiskanen, W.A., and Moritz, H. (1967) Physical Geodesy. Freeman, San Francisco.

Kaula, W.M., Parmenter, M.E., Burkhard, N., and Jackson, D.D. (1975) Application of inversion theory to new satellite systems for determination of the gravity field. Report AFCRL-TR-75-0450.

Knopoff, J., and Jackson, D.D. (1973) The analysis of underdetermined and overdetermined systems. Preprint, Inst. of Geophysics, Univ. of California, Los Angeles.

Krarup, T. (1969) A contribution to the mathematical foundation of physical geodesy. Publ. No. 44, Danish Geodetic Institute, Copenhagen.

Lanczos, C. (1964) Linear Differential Operators. Van Nostrand, Princeton.

Lauritzen, S.L. (1973) The probabilistic background of some statistical methods in physical geodesy. Publ. No. 48, Danish Geodetic Institute, Copenhagen.

Lavrentiev, M.M. (1967) Some Inproperly Posed Problems of Mathematical Physics. Springer, Berlin.

Liebelt, P.B. (1969) An Introduction to Optimal Estimation. Addison-Wesley, Reading, Massachusetts.

Meschkowski, H. (1962) Hilbertsche Räume mit Kernfunktion. Springer, Berlin.

Moritz, H. (1970)Least-squares estimation in physical geodesy. Report No. 130, Dept. of Geodet. Sci., Ohio State Univ.

Moritz, H. (1971) Series solutions of Molodensky's problem. Publ. No. A 70, Deutsche Geodätische Kommission, München.

Moritz, H. (1972) Advanced least-squares methods. Report No. 175, Dept. of Geodet. Sci., Ohio State Univ.

Rao, C.R., and Mitra, S.K. (1971) Generalized Inverse of Matrices and its Applications. Wiley, New York.

Rummel, R. (1976) A model comparison in least squares collocation. Bull. Géod., 50, 181-192.

Schwarz, K.P. (1974) Tesseral harmonic coefficients and station coordinates from satellite observations by collocation. Report No. 217, Dept. of Geodet. Sci., Ohio State Univ.

Schwarz, K.P. (1975) Formulas and Fortran programs to determine zonal coefficients from satellite observations by collocation. Ibid., Report No. 222.

Wiggins, R.A. (1972) The general linear inverse problem: implication of surface waves and free oscillations for earth structure. Reviews of Geophysics and Space Physics, 10(1), 251-285.